Exercise 5

5.1

$y = \mu + \varepsilon$

$p(\varepsilon) = \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau\varepsilon^2}{2}\right)$

(a) $\varepsilon = y - \mu$

$$\Rightarrow p(y_i \mid \mu, \tau) = \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau(y_i-\mu)^2}{2}\right)$$

$\vec{y} = (y_1, \dots, y_N)^T$

$$\overset{iid}{\Rightarrow} p(\vec{y} \mid \mu, \tau) = \left(\sqrt{\frac{\tau}{2\pi}}\right)^N \prod_{i=1}^{N} \exp\left(-\frac{\tau(y_i-\mu)^2}{2}\right)$$

use log to make expression easier

$$\log p(\vec{y} \mid \mu, \tau) = \log\left(\sqrt{\frac{\tau}{2\pi}}^N\right) + \sum_{i=1}^{N} \log\left(\exp\left(-\frac{\tau(y_i-\mu)^2}{2}\right)\right)$$

$$= \underline{\underline{\frac{N}{2}\log(\tau) - \frac{N}{2}\log(2\pi) - \frac{\tau}{2}\sum_{i=1}^{N}(y_i-\mu)^2}} \qquad (4)$$

(b) $\lambda_0, a_0, b_0 \overset{!}{>} 0$

$$p(\vec{y}, \vec{\theta}) = p(\vec{y} \mid \underbrace{\mu, \tau}_{\vec{\theta}}) \cdot p(\vec{\theta})$$

$$\log(p(\vec{y}, \vec{\theta})) = \log(p(\vec{y} \mid \cancel{\mu,\tau})) + \log(p(\vec{\theta}))$$

$$= \quad (4) \quad + \quad ''$$

$$\log(p(\vec{\theta})) = \log\left(\sqrt{\frac{\lambda_0\tau}{2\pi}}\right) + \log\left(\exp\left(-\frac{\lambda_0\tau}{2}(\mu-\mu_0)^2\right)\right) + \log\left(\frac{b_0^{a_0}\tau^{(a_0-1)}}{\Gamma(a_0)}\right) + \log(\exp(-b\tau))$$

$$= \frac{1}{2}\log(\lambda_0\tau) - \frac{1}{2}\log(2\pi) - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 + \log(b_0^{a_0}\tau^{(a_0-1)}) - \log\Gamma(a_0)$$
$$-b\tau$$

$$= \frac{1}{2}\log(\lambda_0\tau) - \frac{1}{2}\log(2\pi) - \frac{\lambda_0\tau}{2}(\mu-\mu_0)^2 + a_0\log(b_0) + (a_0-1)\log(\tau)$$
$$-\log(\Gamma(a_0)) - b\tau$$

$$\Rightarrow \log(p(\vec{y},\theta)) = \left(\frac{N}{2}+\frac{1}{2}+(a_0-1)\right)\log(\tau) \overset{\text{-}}{+} \left(\frac{N}{2}+\frac{1}{2}\right)\log(2\pi) + \frac{1}{2}\log(\lambda_0)$$

$$-\frac{\tau}{2}\left(\lambda_0(\mu-\mu_0)^2 + \sum_{i=1}^{N}(y_i-\mu)^2\right) + a_0\log(b_0) - \log(\Gamma(a_0)) - b\tau$$

$$= \frac{2a_0-1+N}{2}\log(\tau) - \frac{N+1}{2}\log(2\pi) + \frac{1}{2}\log(\lambda_0)$$

$$-\frac{\tau}{2}\left(\lambda_0(\mu-\mu_0)^2 + \sum_{i=1}^{N}(y_i-\mu)^2\right) + a_0\log(b_0) - \log(\Gamma(a_0)) - b_0\tau \quad (*)$$

c) $$F(q) = \int q(\mu)q(\tau)\log\frac{p(\vec{y}\,|\,\mu,\tau)\,p(\vec{\theta})}{q(\mu)\,q(\tau)}\,d\mu\,d\tau$$

$$= \int q(\mu)q(\tau)\log(p(\vec{y}\,|\,\vec{\theta})\,p(\vec{\theta}))\,d\mu d\tau - \int q(\mu)q(\tau)\log(q(\mu)q(\tau))\,d\mu d\tau$$

$$= \text{"} \quad - \int q(\mu)q(\tau)\log(q(\mu))\,d\mu d\tau - \int q(\mu)q(\tau)\log(q(\tau))\,d\mu d\tau$$

$$= \int q(\mu)\int q(\tau)\log(p(\vec{y},\vec{\theta}))\,d\tau d\mu - \int q(\mu)\log(q(\mu))\,d\mu + \text{const}$$

with descr. in exercise

$$= \int q(\mu)\big(\underbrace{E_{q(\tau)}\log(p(\vec{y},\vec{\theta})) + \text{const}}_{=\hat{p}(\vec{y},\vec{\theta})}\big)\,d\mu - \int q(\mu)\log(q(\mu))\,d\mu + \text{const}$$

log rules

$$= \int q(\mu)\log\left(\frac{\hat{p}(\vec{y},\vec{\theta})}{q(\mu)}\right)d\mu + \text{const}$$

$$= -KL(q(\mu)\,|\,\hat{p}(y,\theta)) + \text{const}$$

to maximize $F(q)$ and $KL \overset{!}{\geq} 0 \Rightarrow$

$$q^*(\mu) = \hat{p}(\vec{y},\vec{\theta}) = \exp\left(E_{q(\tau)}(\log(p(\vec{y},\vec{\theta}))) + \text{const}\right)$$

$$\log(q^*(\mu)) = E_{q(\tau)}(\log(p(\vec{y},\vec{\theta}))) + \text{const}$$

all terms without μ are const $(*)$ $$= -\frac{1}{2}E_\tau(\tau)\cdot\left(\sum_{i=1}^{N}(y_i-\mu)^2 + \lambda_0(\mu-\mu_0)^2\right) + \text{const} \checkmark$$

(d) $\log(q^*(\mu)) = -\frac{1}{2} E_\tau\{\tau\} \left(\sum_{i=1}^{N} (y_i - \mu)^2 + \lambda_0 (\mu - \mu_0) \right) + \text{const}$

$= \quad '' \quad \cdot \left(\sum_{i=1}^{N} (y_i^2 - 2y_i \mu + \mu^2) + \lambda_0 (\mu - \mu_0)^2 \right) + \text{const}$

$= \quad '' \quad \cdot \left(\sum_{i=1}^{N} (y_i^2 - 2\mu_i \mu + \mu^2) + \lambda_0 (\mu^2 - 2\mu\mu_0 + \mu_0^2) \right) + \text{const}$

$= \quad '' \quad \cdot \left(\sum_{i=1}^{N} (y_i^2 - 2y_i \mu) + (N+\lambda_0)\mu^2 - 2\lambda_0 \mu\mu_0 + \underbrace{\lambda_0 \mu_0^2}_{= \text{const}} \right) + \text{const}$

$= -\frac{1\bar{\tau}(N+\lambda_0)}{2}\mu^2 + \bar{\tau}\lambda_0 \mu\mu_0 - \frac{1}{2}\bar{\tau} \left(\sum_{i=1}^{N} (y_i^2 - 2y_i\mu) + \text{const} \right) + \text{const}$

$= \quad '' \quad + \quad '' \quad - ('') \cdot \left(\sum_{i=1}^{N} y_i^2 - \sum_{i=1}^{N} 2y_i\mu \right) + \text{const}$

$= \quad '' \quad + \quad '' \quad - '' \cdot \left(\sum_{i=1}^{N} y_i^2 - 2\mu \sum_{i=1}^{N} y_i \right) + \text{const}$

$= -\frac{\bar{\tau}(N+\lambda_0)}{2}\mu^2 + \bar{\tau}\lambda_0\mu\mu_0 - \frac{1}{2}\bar{\tau} \left(\underbrace{\sum_{i=1}^{N} y_i^2}_{= \text{const}} - 2\mu N \bar{y} \right) + \text{const}$

$= -\frac{\bar{\tau}(N+\lambda_0)}{2}\mu^2 + \bar{\tau}\lambda_0\mu\mu_0 + \bar{\tau}N\bar{y}\mu + \text{const}$

$= -\frac{\bar{\tau}(N+\lambda_0)}{2}\mu^2 + \bar{\tau}\bar{\mu}(\lambda_0\mu_0 + N\bar{y}) + \text{const}$

(e) log of a gaussian :

$\log(\mathcal{N}(\mu | m, s^{-2})) = \log\left(\frac{1}{\sqrt{2\pi s^{-2}}}\right) + \log\left(\exp\left(-\frac{(\mu - m)^2}{2s^2}\right)\right)$

$= \underbrace{\log(1)}_{=0} - \frac{1}{2}\log(2\pi s^{-2}) + -\frac{(\mu - m)^2}{2s^2}$

$= \underbrace{-\frac{1}{2}\log(2\pi)}_{\text{const}} - \underbrace{\log(s^{-1})}_{\text{const}} - \underbrace{\frac{(\mu - m)^2}{2s^{-2}}}_{= -\frac{\mu^2}{2s^2} + \frac{2m\mu}{2s^2} - \underbrace{\frac{m^2}{s^2}}_{= \text{const}}}$

$-\frac{\bar{\tau}(N+\lambda_0)}{2}\mu^2 + \bar{\tau}\mu(\lambda_0\mu_0 + N\bar{y}) \stackrel{!}{=} -\frac{\mu^2 - 2m\mu}{2s^{-2}}$

$= -\frac{\mu^2}{2s^2} + \frac{2m\mu}{2s^2}$

first term with μ^2

$$-\frac{1}{2s^2} = -\frac{\bar{\tau}(N+\lambda_0)}{2}$$

$$\frac{1}{s^2} = \bar{\tau}(N+\lambda_0) \Rightarrow s^2 = \frac{1}{\bar{\tau}(N+\lambda_0)}$$

2nd term with μ

$$\bar{\tau}\mu\lambda_0\mu_0 + \bar{\tau}N\bar{y}\mu = \frac{2m\mu}{2s^2}$$

$$\text{"} + \text{"} = \frac{2m\mu \cdot \bar{\tau}(N+\lambda_0)}{2}$$

$$2\mu(\lambda_0\mu_0 + N\bar{y}) = 2m\mu(N+\lambda_0)$$

$$\frac{2\mu(\lambda_0\mu_0 + N\bar{y})}{N+\lambda_0} = 2m\mu$$

$$\frac{\lambda_0\mu_0 + N\bar{y}}{N+\lambda_0} = m$$

(f) same as c) but for $q(\tau)$ see derivations there

$$\log(q(\tau)) = \mathbb{E}_{q(\mu)}(\log(p(\vec{y},\vec{\theta})) + \text{const}$$

$$= -\frac{\tau}{2}E_\mu\left\{\sum_{i=1}^{N}(y_i-\mu)^2 + \lambda_0(\mu-\mu_0)^2\right\} - b_0\tau + \frac{2a_0-1+N}{2}\log(\tau)$$

$$+\text{const}$$

plug in log joint and put everything not dep on τ in const

& use

$$E_x\{f(x)\} = \int f(x)q(x)dx$$

(9) $E_\mu\{(x-\mu)^2\} = (x-m)^2 + s^2$

$$E_\mu\left\{\sum_{i=1}^{N}\left(y_i^2 - 2y_i\mu + \mu^2\right) + \lambda_0\mu^2 - 2\lambda_0\mu\mu_0 + \underbrace{\lambda_0\mu_0^2}_{\text{const}}\right\}$$

$$E_\mu\left\{\underbrace{\sum_{i=1}^{N} y_i^2}_{\text{const}} - 2\mu\sum_{i=1}^{N} y_i + (\lambda_0 + N)\mu^2 - 2\lambda_0\mu_0\mu\right\}$$

$$E_\mu\left\{-2\mu\sum_{i=1}^{N} y_i + (\lambda_0 + N)\mu^2 - 2\mu\lambda_0\mu_0 + \text{const}\right\}$$

$$E_\mu\left\{-2\mu(\lambda_0\mu_0 + N\bar{y}) + (\lambda_0 + N)\mu^2 + \text{const})\right\} \stackrel{!}{=} E_\mu\left\{x^2 - 2\mu x + \mu^2\right\}$$

$$x = \frac{\lambda_0\mu_0 + N\bar{y}}{\lambda_0 + N}$$

(h) $\mathrm{Gam}(\tau \mid a, b) = \dfrac{b^{a}\tau^{a-1}}{\Gamma(a)} \exp(-b\tau)$

$$\log(\mathrm{Gam}(\tau \mid a,b)) = \underbrace{a\log(b)}_{\text{const}} + (a-1)\log(\tau) - \underbrace{\log(\Gamma(a))}_{\text{const}} - b\tau$$

look at terms with $\log(\tau)$

$$a_0 + \frac{N+1}{2} - 1 \overset{!}{=} (a-1)$$

$$\boxed{a = a_0 + \frac{N+1}{2}}$$

look at terms wit τ

$$\tau\left(-\frac{1}{2}\left(\sum_{n=1}^{N}(y_n - m)^2 + \lambda_0(\mu_0 - m)^2 + (N+\lambda_0)s^2\right) - b_0\right) = -b\tau$$

$$\Rightarrow \boxed{b = b_0 + \frac{1}{2}\left(\sum_{n=1}^{N}(y_n-m)^2 + \lambda_0(\mu_0 - m)^2 + (N+\lambda_0)s^2\right)}$$

(i) see derivation in c

$$\mathcal{F}(q) = \int q_1(\mu)\, q_2(\tau) \log\left(\frac{p(\vec{y}, \mu, \tau)}{q_1(\mu)\, q_2(\tau)}\right) d\mu\, d\tau$$

$$= \mathbb{E}_q\left(\log \frac{p(\vec{y} \mid \vec{\theta})\, p(\vec{\theta})}{q(\vec{\theta})}\right)$$

$$= -KL\left(q_1(\mu) \,\|\, \underbrace{\hat{p}(\vec{y}, \mu, \tau)}_{\text{const}}\right)$$

$$= \int \mathcal{N}(\mu \mid m, s^2)\, \mathrm{Gam}(\tau \mid a, b) \log\left(\frac{\overbrace{p(\vec{y} \mid \mu, \tau)\, p(\mu, \tau)}^{p(\vec{y}, \mu, \tau)}}{\mathcal{N}(\mu \mid m, s^2)\, \mathrm{Gam}(\tau \mid a, b)}\right) d\mu\, d\tau)$$

$$= -\frac{1}{2}\mathbb{E}_\tau\{\tau\} \cdot \left(\mathbb{E}_\mu\left\{\sum_{n=1}^{N}(y_n - \mu)^2 + \lambda_0(\mu - \mu_0)^2\right\} + 2b_0\right) + \mathbb{E}_\tau\{\log(\tau)\}\left(a_0 + \frac{N+1}{2} - 1\right)$$

take everything from $\log q^*(\tau)$ with τ-term

everything with $\log(\tau)$

$+ \dots$

$$\text{i}\quad \mathcal{F}(q) = \int q(\mu)\, q(\tau)\, \log\left(\frac{p(\vec{g},\tau,\mu)}{q(\mu)q(\tau)}\right) d\mu\, d\tau$$

$$= \int q(\mu) q(\tau) \log(\vec{p},\tau,\mu)\, d\mu\, d\tau$$

$$- \int q(\mu) q(\tau) \log(q(\mu) q(\tau))\, d\mu\, d\tau$$

$$= \left.\int q(\mu) q(\tau) \log(\vec{p},\tau,\mu)\, d\mu\, d\tau \right\} \text{1}$$

$$- \int q(\mu)\, q(\tau) \log(q(\mu))\, d\mu\, d\tau$$

$$- \int q(\mu)\, q(\tau) \log(q(\tau))\, d\mu\, d\tau$$

$$= \text{1} \quad - \int q(\tau) \underbrace{\int q(\mu) \log(q(\mu))\, d\mu}_{= E_\mu\{\log(q(\mu))\}}\, d\tau$$

$$- \int q(\mu) \underbrace{\int q(\tau) \log(q(\tau))\, d\tau}_{E_\tau\{\log(q(\tau))\}}\, d\mu$$

$$= \text{1} \; - \underbrace{\int q(\tau)\, d\tau}_{=1} \cdot E_\mu\{\log q(\mu)\} - \underbrace{\int q(\mu)\, d\mu}_{=1} \cdot E_\tau(\log q(\tau))$$

$$= \int q(\mu) q(\tau) \log(\vec{p},\tau,\mu)\, d\mu\, d\tau \; \underbrace{- E_\mu\{\log q(\mu)\} - E_\tau((\log q(\tau))}_{\text{2}}$$

$$= \int q(\mu) \int q(\tau) \log(\vec{p},\tau,\mu)\, d\tau\, d\mu + \text{2}$$

$$= \int q(\mu)\, E_\tau\{\log(\vec{p},\tau,\mu)\}\, d\mu + \text{2}$$

$$= E_\mu\{E_\tau\{\log(\vec{p},\tau,\mu)\}\} + \text{2}$$

$$= -\frac{1}{2} E_\tau\{\tau\} \cdot \left(E_\mu\left\{\sum_{n=1}^{N}(g_n-\mu)^2 + \lambda(\mu-\mu_0)\right\} + 2b_0\right) + E_\tau\{\log(\tau)\} \cdot \left(a_0 + \frac{N+1}{2} - 1\right)$$

$$+ \frac{1}{2}\log(\lambda) - \frac{N+1}{2}\log(2\pi) + a_0 \log(b_0) - \log(\Gamma(a_0)) + \text{2}$$